

WHAT IS CLAIMED IS:

1. A method of determining user interactions comprising the steps of:
determining speech information;
determining discourse functions and prosodic features in the speech
5 information;
determining a predictive interaction model; and
determining an interaction turn based on the predictive interaction
model and the determined discourse functions and prosodic features.
2. The method of claim 1, in which the discourse functions are
10 determined from automatically recognized speech information.
3. The method of claim 1, in which the discourse functions are
determined based on a theory of discourse analysis.
4. The method of claim 3, in which the theory of discourse analysis is
at least one of: the Linguistic Discourse Model, the Unified Linguistic
15 Discourse Model, Rhetorical Structures Theory, Discourse Structure Theory
and Structured Discourse Representation Theory.
5. The method of claim 1, further comprising the step of scheduling an
interaction event based on the turn prediction.
6. The method of claim 1, in which the prosodic features include at least
20 one of: a silence preceding a discourse functions; a silence following a
discourse function; rate of speech; pitch frequency; changes in pitch frequency
and volume.
7. A method of determining a predictive interaction model comprising the
steps of:
25 determining a training corpus of turn annotated speech information;
determining discourse functions and prosodic features associated with
the turn information; and
determining a predictive interaction model based on the discourse
functions, the prosodic features and the turn information.
- 30 8. The method of claim 7, in which the predictive interaction model is
determined based on at least one of machine learning, decision tree, Naïve
Bayes, rules and statistics.

9. The method of claim 7, in which the discourse function determined based on a theory of discourse analysis.

10. A system for determining interactions comprising:
an input/output circuit for retrieving recognized speech and prosodic features;

a processor that determines speech information; and discourse functions and prosodic features in the speech information; determines a predictive interaction model; and determines an interaction turn based on the predictive interaction model and the discourse functions and prosodic features.

11. The system of claim 10, in which the discourse functions are determined from automatically recognized speech information.

12. The system of claim 10, in which the discourse functions are determined based on a theory of discourse analysis.

13. The system of claim 12, in which the theory of discourse analysis is at least one of: the Linguistic Discourse Model, the Unified Linguistic Discourse Model, Rhetorical Structures Theory, Discourse Structure Theory and Structured Discourse Representation Theory.

14. The system of claim 10, in which the processor also schedules an interaction event based on the turn prediction.

15. The system of claim 10, in which the prosodic features include at least one of: a silence preceding a discourse functions; a silence following a discourse function; rate of speech; pitch frequency; changes in pitch frequency and volume.

16. Computer readable storage medium comprising: computer readable program code embodied on the computer readable storage medium, the computer readable program code usable to program a computer to determine interactions comprising the steps of:

determining speech information;

determining discourse functions and prosodic features in the speech information;

determining a predictive interaction model; and

determining an interaction turn based on the predictive interaction model and the determined discourse functions and prosodic features.

17. The method of claim 1, in which the speech information is at least one of: verbal natural language information and non-verbal natural language information.

5 18. The method of claim 17, in which the non-verbal information is at least one of: sign language gestures, pen gestures, hand gestures, body gestures and facial gestures.

19. The method of claim 7, in which the speech information is at least one of: verbal information natural language and non-verbal natural language information.

10 20. The method of claim 18, in which the prosodic features include at least one of: facial expressions, gesture velocity, and gesture force.